THE UNIVERSITY OF ALBERTA

EFFECTS OF VERBAL CUES AND CARDINAL NUMBER ON CLASS INCLUSION ERRORS

by

© DIANE MARGARET BORWICK LOVE

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE

OF MASTER OF ARTS

DEPARTMENT OF PSYCHOLOGY

EDMONTON, ALBERTA

FALL, 1976

ABSTRACT

Piaget (1952) discovered that children in the preoperational stage could not solve the class inclusion problem; that is, they do not understand that a subordinate class is included in and thus smaller than its superordinate class. In the present experiment, three factors were examined, which, on the basis of previous research (Winer, 1974; Brainerd and Kaszor, 1974; Gelman, 1972), were expected to affect performance on the class inclusion problem: (1) verbal cueing about the size of classes; (2) equal as opposed to unequal subclass ratio; and (3) small versus large superordinate class size. Analysis of the data revealed no significant main effects. Various explanations for this failure to support the experimental predictions are offered and assessed.

ACKNOWLEDGMENT

I would like to thank Dr. C. Brainerd for the encouragement and advice he offered throughout the preparation of this thesis. I thank also Dr. T. Weckowicz and Dr. L. Stewin for their helpful comments and suggestions.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I

INTRODUCTION

Piaget (1952) discovered that children in the preoperational stage of cognitive development are unable to understand that a subordinate class is included in and thus smaller than its superordinate class. In the original version of the class inclusion problem, the experimental materials consisted of a set of wooden beads of which two were white, while all the others were brown. Surprisingly, even when the subjects understood that the brown beads and the white beads were all at the same time wooden beads, most could not correctly answer the question, "Which would be longer, a necklace made of the brown beads or one made with the wooden beads?" The typical reply was "The one made with the brown beads."

Varying the experimental materials did not change the results. Piaget used, for example, a set of flowers, 20 poppies and two or three bluebells, and asked, "Which will be the bigger bunch, one made with all the flowers, or one made with all the poppies?" Again, his young subjects were unable to arrive at the correct answer.

These first class inclusion experiments were conducted within the context of Piaget's study of number conservation. The common basis underlying both concepts, class inclusion and number conservation, was assumed to be "the additive operation, which brings together the scattered elements into a whole, or divides these wholes into parts." In other words, class inclusion may be viewed as a problem of additive composition, such that two (or

more) subclasses, A and A', are included in a superordinate class, B. Understanding class inclusion then, implies understanding of the following relationships: $A + A' = B$; therefore, $A = B - A'$, and $A < B$. According to Piaget, failure to solve the wooden bead problem is evidence of the child's inability to grasp the relationship $A < B$, that is, that the subordinate class is necessarily smaller than the superordinate class.

Inhelder and Piaget (1964) subsequently confirmed the earlier results. Using sequences of four or five hierarchical classes, such that $A < B < C < D < E$, Inhelder and Piaget asked their subjects a lengthy series of questions designed to probe deeper into the class inclusion problem. In these experiments, stimulus materials were flowers, beads or pictures of animals. Again, children at the preoperational stage, approximately 7 or 8 years of age, could not correctly answer questions such as "Are there more birds or more animals?". Indeed, the animal classes proved most difficult of the three types of stimuli, possibly because the children were less familiar with animals than with flowers or beads: "zoological classes are not very clearly defined for them."

To restate the issue, preoperational children fail to understand that a subordinate class is smaller than its superordinate class. This is often the case even when they know that a member of the subordinate class is at the same time a member of the superordinate class; for example, a bird is also an animal. Why is this so? Inhelder and Piaget postulate that instead of comparing the

subclass, A, with the superordinate class, B, subjects compare A with the complementary subclass, A'. Because A' is smaller than A, they state that there are more A's than B's. In order to arrive at the correct conclusion, it is necessary to hold in mind the identity of the whole, B, and at the same time, separate B into its component parts, A and A'. This is precisely what the preoperational child cannot do. As soon as he or she has separated B into A and A', the identity of B is lost, and thus the comparison becomes A with A' rather than A with B.

Because understanding class inclusion involves an awareness of the reversibility of the relations, $A + A' = B$, and $A = B - A'$, Inhelder and Piaget conclude that class inclusion is a "genuine logical operation". In their words,

> "it is one thing to carry out the union expressed by $A + A' = B$ and quite another to understand that it is logically equivalent to its inverse, $A = B - A'$, which means that the whole, B, retains its identity and that the entire relation can be quantitatively expressed in the form $A < B$. The conservation of the whole and the quantitative comparison of whole and part are the two essential characteristics of genuine class-inclusion;".

Inhelder and Piaget have generated considerable interest in the class inclusion problem. Subsequent researchers (Kofsky, 1966; Wohlwill, 1968; Ahr and Youniss, 1970) have placed the age of solution of the class inclusion problem at anywhere from 6 to 10 years, and numerous variables have been investigated for their effects on children's ability to solve the problem. In general, class inclusion studies conducted during the past 10 years may be classified as dealing either with the effects of training, or with

task variables. Within the latter category, most experiments have been concerned with linguistic and/or perceptual factors that affect task performance. The format of the class inclusion problem has remained basically similar to that used by Piaget in that the subject is required to say which is larger, the subordinate class or the superordinate class.

Among the training studies is that reported by Sheppard (1973) who found that 6-year-old subjects who had received training on the class inclusion problem performed significantly better on post-tests administered up to three or four months later, than those who received no training. However, it should be noted that the effect of training was only great enough to increase the number of correct responses from a mean of $<1/10$ (control group) to a mean of approximately 4--5/10 (experimental group).

In another recent training study, Brainerd (1974) found that 4- and 5-year-old subjects who were trained via verbal feedback on three concepts including class inclusion of length and weight, showed an improvement in performance for all concepts with class inclusion having the smallest improvement.

Using correction training, Ahr and Youniss (1970) also obtained improved performance on the class inclusion task. Their method consisted of informing subjects of all errors made, and requiring them to supply the correct answer. Again, in both this and Brainerd's studies, the effect of training was not dramatic, and it seems reasonable to conclude that training in class inclusion does not produce any very important change in performance. Therefore,

whether they receive training or not, subjects aged less than 8 years simply do not do well on the class inclusion task.

Leaving the issue of training, there is a second group of studies in which various aspects of the class inclusion task, including linguistic and perceptual factors, have been manipulated. To illustrate, Wohlwill (1968) found that presenting the problem in a purely verbal form produced better performance than a combined verbal and pictorial presentation. Specifically, subjects who were shown pictures representing the subclass-superordinate class comparison scored lower on the class inclusion questions than did subjects who did not see any pictures. Wohlwill interpreted these results as being due to a misleading perceptual set present in the pictorial condition that caused subjects to make a comparison of the sizes of the two subclasses rather than the correct comparison of the subordinate and superordinate classes. Presumably, in the purely verbal form of the task, this perceptual set is eliminated, thus accounting for the better performance in this condition.

However, other researchers (Jennings, 1970; Brainerd & Kaszor, 1974) have failed to confirm Wohlwill's results. In fact, Jennings, using Wohlwill's items, found that subjects in grades 1 through 3 performed significantly better in the pictorial condition, which is exactly opposite to Wohlwill's findings; while for kindergartners there was "essentially no difference between the verbal and pictorial conditions". It should be noted that the methodology used by Jennings involved asking subjects to justify their responses and to show their understanding of the questions; subjects who were able

to do so received higher scores than those who could not. Jennings emphasizes the importance of using such a methodology to assess accurately the effect of the main experimental variables.

Further doubt was cast on the validity of Wohlwill's results by Winer (1974) who demonstrated that the verbal facilitation effect was due to different phrasing of the question in the verbal condition versus pictorial condition, rather than to the presence or absence of pictures as Wohlwill had stated. In the pictorial condition of Wohlwill's experiment, subjects were asked simply, "Are there more dogs or more animals?" In the purely verbal presentation, the number of objects in each subclass was mentioned; in other words, the question was "If I had seven dogs and three horses, would I have more dogs or more animals?" The additional information in the latter version of the question apparently served to remind subjects of the relationship between the superordinate class and its subclasses.

In Winer's experiment, subjects from grades 2 through 4 were asked three sets of questions, which were not accompanied by pictures. Here subjects were informed of the number of objects in each subclass. A second set of questions was similar to Wohlwill's pictorial condition in that subclass size was not mentioned in the questions, which were accompanied by pictures. In the third set of questions, subjects were also shown pictures but in this case, the questions were verbally elaborate, as in the first set. Results indicate that the verbally elaborate questions, whether or not they were accompanied by pictures, were answered correctly more often than those questions,

in the pictorial condition, that made no mention of subclass size. Thus, Winer concludes that it is verbal cueing rather than absence of misleading perceptual cues that is responsible for improved performance.

One of the objectives of the present study was to further analyze the effect of such verbal cueing. To this end, the form of the question regarding the superordinate-subordinate relationship was varied in terms of amount of information provided as to the size of each class. Specifically, there were four levels of information: in the first form of question, class size was not mentioned; in the second only the size of the subordinate classes was given; while in the third, only the size of the superordinate class was provided; and in the fourth, the size of all classes was mentioned. It was predicted that performance would improve as amount of verbal information increased.

A second variable explored in this study was difference in subclass size. Ahr and Youniss (1970) reported that performance improves as the number of objects in each subclass approaches equality. It appeared that when there was a large numerical difference in subclass size, subjects were more likely to compare the two subclasses rather than the subclass and the superordinate class, and therefore respond incorrectly. Using "pets" and "flowers" as the superordinate classes, Ahr and Youniss varied the subclass ratio in each superordinate class from maximum inequality, 8:0, through ratios of 7:1, 6:2, 5:3 to equality, 4:4. Thus, for example, in the task using "pets" as the superordinate class, the size of the

subclasses changed from eight dogs and no cats through to four dogs and four cats. For each of the five subclass ratios, two forms of question were asked: "Are there more pets or more dogs?" and "Are there fewer pets or fewer dogs?"

Results indicated that for children between 6 and 8 years of age, the subclass ratios 4:4 and 5:3 accounted for over half of the correct responses, and furthermore, as the numerical difference between subclasses increased, performance in this age group deteriorated. Ten-year-old subjects performed much better than the younger subjects and "their performance was less sensitive to numerical differences between subclasses." Ahr and Youniss attributed their results to misleading perceptual cues and concomitant question misinterpretation, both of which have stronger effects with increasing disparity of subclass size.

However, Brainerd and Kaszor (1974) recently demonstrated that this effect of numerical disparity in subclass size is an artifact of the questions employed by Ahr and Youniss. As discussed previously, the questions employed for each subclass ratio were of the form "Are there more B or more A?" and "Are there fewer B or fewer A?" (where A denotes the subclass and B, the superordinate class). When these two questions are posed in conjunction with the subclass ratio 4:4, the subject who does not have the class inclusion concept is more likely to respond correctly than when the subclasses are unequal. Ahr and Youniss attributed this tendency to decreased perceptual disparity, but Brainerd and Kaszor explain it differently: the subject who compares A and A' instead of A and B observes that

A is neither more nor fewer than A' when the subclass ratio is 4:4. But because the questions require the subject to choose either A or B as more or fewer, he or she responds randomly, sometimes choosing A, and sometimes B, which is the correct answer. Therefore, by chance, many of the subject's responses will be scored as correct even though he or she has no understanding of class inclusion.

Evidence that this analysis is correct is provided in the results of an experiment conducted by Brainerd and Kaszor. They obtained "virtually perfect" performance when questions of the form "Are there more/fewer A or more/fewer B?" were asked in conjunction with equal subclasses. Furthermore, similarly high performance occurred when questions of the form "Are there the same number of A's as B's?" were posed using unequal subclasses. These results, then, are incompatible with the misinterpretation/perceptual set hypothesis offered by Ahr and Youniss.

Additional disconfirming evidence, also reported by Brainerd and Kaszor, is that when subjects were required to recall the question preceding an incorrect response, the majority either remembered the question correctly or had forgotten it. There were few recall errors of the type predicted by the misinterpretation hypothesis, i. e. not many subjects recalled the question as asking for a comparison of subclasses.

As discussed previously, Ahr and Youniss stated that performance on the class inclusion problem improved as the subclass ratio approached equality. There is some evidence that exactly the opposite is true. In a second experiment, using pictures of

animals (horses and dogs), Brainerd and Kaszor found that "extreme disparity (A = 9/A' = 1) produces more correct judgments than either moderate disparity (A = 7/A' = 3) or equality (A = 5/A' = 5)." The second purpose of the present study was to discover whether or not these results obtained by Brainerd and Kaszor are replicable. The prediction was that there would be more correct responses when subclasses were unequal than when subclasses were equal.

A third objective was to test the hypothesis that children will be better able to solve the class inclusion problem if smaller numbers of objects are involved. Typically class inclusion researchers have used a total of eight to ten objects in the superordinate class. One of the objectives of the present study was to investigate the effect of using a small number of objects--four--as compared to a large number of objects--eight. It was expected that children would perform better when the class size was small.

Although there is no direct evidence pertaining to the issue of class size in the class inclusion problem, studies of number concept development have indicated that children can make judgments of numerosity much more easily when small numbers are involved than when the experimental task involves large numbers. The bulk of the relevant evidence is provided by Gelman (1972) who cites extensive studies that demonstrate that for preschool children, any number greater than three tends to be perceived and understood as simply "many".

Gelman also reports that experiments with school-aged children indicate that first- and second-grade subjects are capable of

accurately estimating the numerosity of up to and including five objects. For larger numbers, they perform with less accuracy. The tasks used in these experiments included estimation of numerosity of arrays of dots or of objects thrown randomly on a table; requiring subjects to draw a certain number of marbles on paper; or requesting subjects to give the experimenter particular numbers of objects.

Two recent studies of the use of various cues in number judgments made by preschoolers indicate that this age group attends to relevant cues more often when numbers are small, i. e. within the range of estimability, than when numbers are large (Lawson, Baron & Siegel, 1974; Smither, Smiley & Rees, 1974). There is reason to suppose, then, that children's performance on the class inclusion problem will be better when the superordinate class is small.

To reiterate, there were three main experimental variables investigated in the present study:

1) Verbal information regarding class size. There were four levels of verbal information, ranging from no information given about the size of subordinate or superordinate classes, to information regarding the number of objects in either subordinate or superordinate classes, to information about the size of all three classes. It was predicted that as amount of verbal information increased, performance would improve. Presumably, telling subjects how many objects were present would remind them of the superordinate-subordinate class relationship, that is, that the subclass contains fewer objects than the superordinate class.

2) The second variable examined was unequal versus equal subclass size. Subclass ratios were 3:1, 6:2 and 2:2, 4:4. The prediction was that unequal subclass size, 3:1 and 6:2, would produce more correct answers.

3) Large versus small superordinate class size was the third variable explored. A total of eight objects were used in the large superordinate class condition and four objects in the small superordinate condition. The latter was predicted to lead to better performance.

A fourth variable was the age of subjects and on the basis of past research, it was predicted that older children would arrive at more correct answers than younger children.

CHAPTER II

METHOD

Subjects were 96 randomly selected children, 32 from each of grades one, two and three of an Edmonton school. I served as experimenter.

According to Klahr and Wallace (1972), "those CI tasks that use meaningless (in the sense of having no semantic hierarchy) objects, . . . best test the formal ability to deal with class inclusion, independently of ability to deal with semantic hierarchies." Consequently, in the present experiment, it was decided that such classes as pets or flowers should not be used. Instead, materials consisted of coloured circles, black and red, drawn on sheets of manila board, and similar to the stimuli used by Brainerd and Kaszor (1974). There were four different sets of circles, representing the two variables unequal versus equal subclass size and small versus large superordinate class size. The sets consisted of: one black and three red circles (small superordinate class, unequal subclasses); two black and two red circles (small superordinate class, equal subclasses); two black and six red circles (large superordinate class, unequal subclasses); four black and four red circles (large superordinate class, equal subclasses). The rationale behind the choice of class size was first, to select one small superordinate class size--four circles--that is within the easy comprehension of young children, and to compare performance in this condition to performance using a large superordinate class--eight circles--which is the usual number of objects employed in

previous class inclusion experiments. The second reason for choosing these particular class sizes was that the subclass ratios, 3:1 and 6:2, and 2:2 and 4:4 maintain proportionate subclass sizes in both the small and large superordinate classes.

There is some evidence that an intermingled arrangement of stimuli results in better performance than a spatially segregated arrangement (Wohlwill, 1968). In the present study, therefore, the black and red circles were mixed together within each set.

Each subject was assigned to one of four verbal information conditions: (a) No verbal information provided concerning number of objects in the subclasses or the superordinate class; (b) verbal information given regarding number of objects in the subclasses only; (c) verbal information provided as to number of objects in the superordinate class only; and (d) verbal information given concerning number of objects in both the subclasses and the superordinate class. Assignment to conditions was random.

Every subject was asked two questions about each of the four sets of circles. As illustrated in Table 1, the questions depended upon the verbal information condition to which the subject was assigned, and also upon whether subclasses were equal or unequal in size. For example, in condition "a" and when the two subclasses were unequal, the questions were, "Here is a picture of some dots. Some dots are red and some are black. Are there more red dots than there are dots? Are there fewer red dots than there are dots?"

For the stimuli in which the two subclasses were of equal size (and in condition "a"), the questions were phrased as "Here is

TABLE 1

Class Inclusion Questions

Subclass ratio	
	Information Condition "a"
3:1	Some dots are red and some are black.
	Q. 1: Are there more red dots than there are dots? Q. 2: Are there fewer red dots than there are dots?
2:2	Some dots are red and some are black.
	Q. 1: Are there the same number of red dots as there are dots? Q. 2: Are there the same number of dots as there are red dots?
6:2	Some dots are red and some are black.
	Q. 1: Are there more red dots than there are dots? Q. 2: Are there fewer red dots than there are dots?
4:4	Some dots are red and some are black.
	Q. 1: Are there the same number of red dots as there are dots? Q. 2: Are there the same number of dots as there are red dots?
	Information Condition "b"
3:1	There are 3 red dots and 1 black dot.
	Q. 1: Are there more red dots than there are dots? Q. 2: Are there fewer red dots than there are dots?
2:2	There are 2 red dots and 2 black dots.
	Q. 1: Are there the same number of red dots as there are dots? Q. 2: Are there the same number of dots as there are red dots?

(TABLE 1 continued)

6:2	There are 6 red dots and 2 black dots.
	Q. 1: Are there more red dots than there are dots? Q. 2: Are there fewer red dots than there are dots?
4:4	There are 4 red dots and 4 black dots.
	Q. 1: Are there the same number of red dots as there are dots? Q. 2: Are there the same number of dots as there are red dots?
	Information condition "c"
3:1	There are 4 dots altogether.
	Q. 1: Are there more red dots than there are dots? Q. 2: Are there fewer red dots than there are dots?
2:2	There are 4 dots altogether.
	Q. 1: Are there the same number of red dots as there are dots? Q. 2: Are there the same number of dots as there are red dots?
6:2	There are 8 dots altogether.
	Q. 1: Are there more red dots than there are dots? Q. 2: Are there fewer red dots than there are dots?
4:4	There are 8 dots altogether.
	Q. 1: Are there the same number of red dots as there are dots? Q. 2: Are there the same number of dots as there are red dots?
	Information Condition "d"
3:1	There are 3 red dots and 1 black dot. Altogether there are 4 dots.
	Q. 1: Are there more red dots than there are dots? Q. 2: Are there fewer red dots than there are dots?

(TABLE 1 continued)

2:2	There are 2 red dots and 2 black dots. Altogether there are 4 dots.
	Q. 1: Are there the same number of red dots as there are dots?
	Q. 2: Are there the same number of dots as there are red dots?
6:2	There are 6 red dots and 2 black dots. Altogether there are 8 dots.
	Q. 1: Are there more red dots than there are dots?
	Q. 2: Are there fewer red dots than there are dots?
4:4	There are 4 red dots and 4 black dots. Altogether there are 8 dots.
	Q. 1: Are there the same number of red dots as there are dots?
	Q. 2: Are there the same number of dots as there are red dots?

a picture of some dots. Some dots are red and some are black. Are there the same number of red dots as there are dots? Are there the same number of dots as there are red dots?"

These particular versions of questions avoid the type of questioning error made by Ahr and Youniss (1970) and discussed by Brainerd and Kaszor (1974), and yet both contain the same amount of information.

In conditions "b", "c", and "d", respectively, mention was made of (1) the actual number of circles of each colour, that is, in each subclass; (2) the total number of circles, that is, in the superordinate class; and (3) the actual number of circles of each colour and the total number of circles.

To control for any possible effects of order of presentation, the four different stimuli were presented in four sequences, such that no stimulus ever succeeded another more than once. The orders of presentation were: 1, 2, 3, 4; 2, 4, 1, 3; 3, 1, 4, 2; and 4, 3, 2, 1. Subclass ratio 3:1 was designated as stimulus number 1, subclass ratio 2:2 as stimulus number 2, subclass ratio 6:2 as stimulus number 3, and subclass ratio 4:4 as stimulus number 4.

Subjects' responses were recorded on paper by the experimenter. A correct answer was scored as 1, and an incorrect answer as 0. Because two questions were posed for each stimulus, the maximum possible score for each stimulus was 2; the maximum possible over all stimuli was 8.

The experiment was a 3 x 4 x 4 x 2 x 2 x 2 split plot design, with the 3 levels of age, the 4 levels of information, and the 4 levels of order assigned between subjects, and the 2 levels of subclass ratios and the 2 levels of superordinate class size assigned within subjects (repeated measures). Data were analyzed using a mixed model analysis of variance.

CHAPTER III

RESULTS

A summary of the data is presented in Table 2. The mean score, out of 8, is given for each of the 4 main variables: amount of information, size of superordinate class, subclass ratio, and educational grade (age). The means for each level of each of the four main variables are presented in Table 3.

TABLE 2

Mean Numbers of Correct Answers

	Small Superordinate Class		Large Superordinate Class	
Verbal Information	Unequal Subclasses	Equal Subclasses	Unequal Subclasses	Equal Subclasses
		Grade 1		
None	1.50	3.00	1.00	5.50
A/A' only	1.00	0.50	2.00	1.00
B only	3.00	2.00	2.50	2.00
A/A' and B	2.50	2.50	2.00	1.00
		Grade 2		
None	2.50	2.50	1.50	3.50
A/A' only	2.50	2.50	1.00	3.00
B only	2.00	5.00	4.00	4.00
A/A' and B	6.00	4.00	5.00	5.00
		Grade 3		
None	2.50	3.00	2.00	3.50
A/A' only	5.00	2.00	4.00	2.50
B only	2.00	4.00	1.50	3.50
A/A' and B	3.50	3.50	3.00	3.00

TABLE 3

Mean Number of Correct Answers for Main Variables

Verbal Information			
None	A/A' only	B only	A/A' and B
2.67	2.25	3.04	3.38

Superordinate Class Size	
Small	Large
2.90	2.77

Subclass Ratio	
Unequal	Equal
2.71	2.96

Grade		
One	Two	Three
2.06	3.40	3.03

An analysis of variance revealed no significant main effects. Further analysis, employing Tukey's test for pairwise comparisons (Kirk, 1968), confirmed the preliminary analysis.

Three interaction effects attained statistical significance: the interaction of information by subclass ratio ($F = 9.48$, $df = 3, 40$, $p < .01$), the interaction of information by subclass ratio by grade ($F = 5.41$, $df = 6, 40$, $p < .01$), and the interaction of subclass ratio by grade by order of presentation ($F = 5.79$, $df = 6, 40$, $p < .01$). These interactions were further analyzed using Scheffe's S method (Kirk, 1968). Within the I x A interaction, when the sub-

class ratio was unequal, providing subjects with no information produced significantly poorer performance than when they were given any amount of information about class size. However, when the subclass ratio was equal, providing information about the size of subclasses gave rise to significantly poorer scores than any of the other information conditions.

Turning to the I x A x G interaction, scores were significantly higher in the equal subclass ratio condition than in the unequal ratio condition for grade one subjects in the "no information" condition, and for grade three subjects in the information condition in which the size of the superordinate class was given. However, when grade three subjects were given information about the size of subclasses, they scored higher when the subclass ratio was unequal than when it was equal.

Comparisons of means within the A x G x O interaction revealed that performance in the equal subclass ratio condition was significantly better than that in the unequal ratio condition when grade one subjects received stimulus presentation order number 2, and also when grade two subjects received stimulus order number 1.

An analysis of variance summary for main effects and first and second order interactions is presented in Table 4. Third and fourth order interactions are not included here because of the very small F ratios obtained, and also because these interactions are of no theoretical importance.

TABLE 4

Analysis of Variance Summary

Source	SS	df	MS	F
Information	4.22	3	1.40	.91
Grade	7.69	2	3.84	2.47
Order	3.38	3	1.13	.72
I x G	8.66	6	1.44	.93
I x O	24.48	9	2.72	1.75
G x O	18.90	6	3.15	2.02
I x G x O	31.25	18	1.74	1.12
Subjects (IGO)	74.75	48	1.56	
Subclass Ratio, A	.38	1	.38	1.95
I x A	5.48	3	1.83	9.48*
G x A	.02	2	.01	.04
O x A	1.92	3	.64	3.32
I x G x A	6.26	6	1.04	5.41*
I x O x A	2.48	9	.28	1.43
G x O x A	6.69	6	1.12	5.79*
Subjects x A (IGO)	9.28	48	.19	
Super. Class, B	.94	1	.94	.49
I x B	.68	3	.23	1.17
G x B	.20	2	.10	.53
O x B	.20	3	.07	.34
I x G x B	.78	6	.13	.67
I x O x B	3.28	9	.36	1.89
G x O x B	1.00	6	.17	.87
Subjects x B (IGO)	9.24	48	.19	
A x B	.84	1	.84	3.77
I x A x B	1.05	3	.35	1.57
G x A x B	.11	2	.06	.24
O x A x B	1.36	3	.45	2.03
Subjects x A x B (IGO)	10.76	48	.22	

*$p < .01$

CHAPTER IV

DISCUSSION

The focus of the present experiment was the development of the concept of class inclusion, which may be described as an understanding that a subordinate class is included in and thus smaller than its superordinate class. Recent research into the class inclusion problem has dealt with several different variables, which may be categorized as dealing basically with task variables (for example, perceptual set) or with training of the concept.

The present study was intended to further investigate two experimental effects uncovered by previous research: verbal cueing about the size of classes and unequal versus equal subclass ratios. The first of these, verbal cueing, was first suggested by Winer (1974), as an explanation of Wohlwill's (1968) "perceptual-set interpretation" of children's failure on the class inclusion problem. Wohlwill had reported that subjects who were shown pictures representing the class inclusion situation did not perform as well as subjects who were presented with the problem in a purely verbal form. However, Winer was subsequently able to demonstrate that this effect is attributable to verbal cueing about the size of subclasses rather than to the absence of "misleading" perceptual cues. In the present study, four different conditions of verbal cueing were investigated: (1) no information about any class; (2) information about the size of the subclasses; (3) information about the size of the superordinate class; and (4) information about both the superordinate and subordinate classes.

An effect with respect to subclass ratio was first reported by Ahr and Youniss (1970), who found that as the ratio approached equality, performance improved. However, Brainerd and Kaszor (1974) showed that these results were due to an artifact of the questioning method. Furthermore, they suggested that there is some evidence to the contrary; that is, that performance is better when subclasses are unequal. In the present study, there were two unequal subclass ratios (3:1 and 6:2) and two equal subclass ratios (2:2 and 4:4).

A third variable, absolute size of superordinate class, has not been previously investigated. However, other research indirectly related to the class inclusion problem served as the basis for prediction. Much of this research has been summarized by Gelman (1972), who reports that young children have difficulty comprehending numerosity beyond a small number of objects. Other research has shown that preschoolers making numerosity judgments attend to relevant cues more often when numbers are small (Lawson, Baron & Siegel, 1974; Smither, Smiley & Rees, 1974). Superordinate classes of four and eight objects were employed in the present experiment.

With regard to these three main variables, it was predicted that performance on the class inclusion problem would improve in conditions of (1) increasing amount of verbal information about the size of classes; (2) unequal subclass ratios; and (3) small superordinate class.

Analysis of the data provided virtually no support for any of these predictions; nor were there any marked trends in directions other than those predicted. The analysis of variance revealed no

significant main effects, and this preliminary analysis was confirmed by post hoc comparisons.

With regard to the first prediction, that performance would improve as verbal cueing increased, post hoc tests revealed only one comparison that could, perhaps, be considered to be supportive evidence; this comparison was within the I x A (Information x Subclass Ratio) interaction, which was significant at $p < .01$. When the subclass ratio was unequal, subjects in the "no information" condition scored significantly lower on the class inclusion questions than did subjects in the three other information conditions. In other words, when there was no mention made of the size of any class, and when the subclasses were unequal, performance was lower than when information was provided about the size of the subclasses or of the superordinate class, or both. Winer (1974) found that verbal cueing about the size of subclasses improved performance, and in his experiment, only unequal subclasses were employed. Thus the effect found in the present study is in accordance with Winer's research.

However, when the subclass ratio was equal, different results were obtained. In this case, subjects who were told the size of the subclasses scored significantly lower than subjects in the other conditions, including the condition of no information about any of the classes. This is a rather surprising finding. It seems that telling subjects how many objects were in each subclass hindered their performance when the subclasses were equal, but did not do so

when the subclasses were unequal. This result may be explained in terms of the question format employed.

When subclasses were equal, the two questions were, "Are there the same number of red dots as there are dots?" and "Are there the same number of dots as there are red dots?" For the unequal subclass ratios, on the other hand, the questions were, "Are there more red dots than there are dots?" and "Are there fewer red dots than there are dots?" Note that for the equal subclasses, the questions are merely different phrasing of the same thing. This is not true of the questions posed in conjunction with the unequal subclasses; here the questions represent different alternatives.

If, in some way, cueing on the size of subclasses was misleading to subjects who were set to make a comparison of the subclasses (rather than the subordinate/superordinate comparison), the questions posed in conjunction with the equal subclasses could lead to a higher probability of incorrect response on the second question, because it would be judged to be the same as the first and thus requiring the same response. In contrast, the questions posed in conjunction with the unequal subclasses were different alternatives. In this case, children would, perhaps, consider the second question separately from the first. That is, they would think about the response to each question separately. Thus, there would be a higher probability of responding correctly to the unequal subclass questions.

To reiterate, the process may be something like the following:

1) Information about the number of dots in each subclass may be a potentially misleading cue in that the question immediately following the cue (for both equal and unequal subclasses) describes a situation ("same number of red dots as dots" or "more red dots than dots") which would match the comparison of red dots with black dots; thus the child who is set to make the subclass comparison will answer the first question incorrectly.

2) When the subclasses are equal, the second question is judged to be simply a re-phrasing of the first. Thus the child is more likely to respond with the same answer for both questions.

3) When the subclasses are unequal, the two questions are judged to be two different alternatives and thus the child is more likely to respond differently, i . e. to ponder both questions rather than only the first. Therefore, the probability of at least one correct response is higher in this case.

An alternative explanation for the obtained interaction of information with subclass ratio is that when the child is given information about the size of subclasses, this information is more likely to be confusing when subclasses are equal than when they are unequal. That is, the question format may not be a relevant factor here; rather, the child who does not have a clear understanding of the class inclusion concept may find subclass information a hindrance when subclasses are equal but not when they are unequal. This would imply that class inclusion is more easily comprehended when subclasses are unequal than when they are equal. Such an interpretation is consistent with the results obtained by Brainerd

and Kaszor (1974), who found that "extreme A/A' disparity improves judgments slightly."

Subclass ratio was the second main experimental variable and the prediction was that unequal subclasses would lead to better performance than equal subclasses. This prediction was based on the aforementioned experiment conducted by Brainerd and Kaszor. Because there was no significant main effect, it must be concluded that there was, in general, no support for this prediction. In fact, a comparison of the means for unequal versus equal subclass ratios at the first level of verbal information (no information about class size) revealed that in this condition, performance was significantly better when the subclasses were equal. However, because there was no overall dramatic improvement, the issue must be considered unsettled as yet.

Relevant to a discussion of subclass ratio are the several significant comparisons uncovered, by post hoc tests, within the two significant three-way interactions. Performance on the questions posed in conjunction with equal subclass ratios was significantly better than that associated with unequal subclass ratios when:

1) subjects in grade one received no information about any of the classes;

2) subjects in grade one received stimulus presentation order number 2;

3) subjects in grade two received stimulus presentation order number 1; and

4) subjects in grade three received information about the size of the superordinate class.

In contrast to these four comparisons, a fifth significant post hoc comparison, within the I x A x G interaction, revealed that in one case performance was better when subclasses were unequal than when they were equal. This was true of grade three subjects who received information about subclass size.

Because of the large number (33) of F ratios computed, and because there is no theoretical basis for expecting such results, these statistically significant interactions are best interpreted as chance fluctuations. In other words, these results contribute nothing to an understanding of the issue of subclass ratios in the development of the class inclusion concept.

Turning to the third main prediction, it was hypothesized that because young children have a better understanding of number when only a few objects are involved, use of a small superordinate class would lead to higher scores than those obtained when the superordinate class was large. There was no significant main effect for this variable, nor did it participate in any significant interactions. Thus results failed to provide any support for the prediction.

A fourth variable, age (educational grade), was also included in the present study. It was predicted that older children would perform significantly better on the class inclusion problem than younger children. However, the mean scores of the grades one, two, and three children who participated in the study were not

significantly different. This is contrary to the findings of previous research, which shows that usually older children do perform at a higher level than younger ones. One possible explanation for the lack of significant differences is that the school from which subjects were drawn was situated in a lower to lower-middle class neighbourhood. Because children from this social class often fall behind their middle to upper class peers on many measures of scholastic achievement, it may be that this would contribute to the low scores of all three grades.

A more likely explanation, or one that would account for a greater portion of performance, is that most children are not capable of grasping the class inclusion relation until they are at least 10 years old. Kofsky (1966), for example, in an extensive study of classificatory development, using subjects of above average intelligence, found that even among the 9-year-old children, performance on the class inclusion problem was not consistently high. In addition, Ahr and Youniss (1970) reported that subjects aged 6 and 8 years did not perform as well as subjects aged 10 years; and even the majority of 10-year-olds did not attain perfect or even near-perfect scores.

To summarize the discussion to this point, it was predicted that children's performance on the class inclusion problem would improve as a function of: increasing amount of verbal information, unequal rather than equal subclass ratios, small rather than large superordinate class, and increasing age of subjects. None of these predictions were supported by the data.

Several different explanations could be advanced to account for the lack of confirming evidence for these predictions. Contributing factors may be grouped into three general categories: (1) experimental design, (2) task-related variables, and (3) subject characteristics. Note that these factors should not be considered to be mutually exclusive.

The present experiment employed a split-plot design with information level, age and order assigned between subjects, and superordinate class size and subclass ratio assigned within subjects (repeated measures). In a split-plot design with repeated measures, the main effects of treatments that are assigned between subjects are completely confounded with differences between sets of subjects. Therefore, tests on between-subjects factors are much less powerful than tests on within-subjects factors, which are free of such confounding. This phenomenon may be partially responsible for the lack of significant main effects for the information and age variables. Because previous research has indicated the existence of these effects, the plausibility of this explanation seems good. It would be interesting to repeat the experiment with the information variable employed as a repeated measure.

Another possibility to be considered is that the experimental task was in some way inappropriate. It might be argued, for example, that the construction of the superordinate and subordinate classes--the class of "dots" composed of "red dots" and "black dots" --is artificial and less familiar to children than, say, the type of hierarchical classes of animals or flowers that are usually

employed in class inclusion experiments. On the other hand, as mentioned previously, classes which are not involved in a semantic hierarchy are probably more appropriate for testing children's understanding of class inclusion than are the more "meaningful" classes such as animals.

A second task-related issue that may be relevant to performance is that of language: to what extent does the choice of questions, for example, affect the child's ability to deal with the class inclusion problem? This of course, is not an issue peculiar to this particular experiment, but one which plays a role in any research employing language. As discussed previously, other research has examined the effect of verbal cueing (Winer, 1974) and some scholars have proposed a linguistic explanation--the misinterpretation hypothesis--for children's failure on the class inclusion problem (Ahr and Youniss, 1970). However, there has been no systematic assessment of psycholinguistic factors in this area, and it is thus impossible to evaluate the effect of any particular version of class inclusion questioning. Nonetheless, it is possible that the particular questions employed in the present study may have adversely affected performance.

The third and probably most important category of explanation for the results obtained in the present study is that of subject characteristics. One such possibility is that children in the age group tested--approximately 7 to 9 years--do not pay close attention to the wording of the class inclusion questions. It was my impression, as experimenter, that this is indeed the case for at

least some subjects. However, attentional variables can probably be ruled out as a major cause of poor performance because of the fact that training on the class inclusion problem does not produce large increments in scores (Brainerd, 1974; Sheppard, 1973). Presumably, in a training situation, the subject's attention is deliberately focussed, by the experimenter, on the problem to be trained. Furthermore, Brainerd and Kaszor (1974) reported that the majority of their subjects were able to recall the questions, which indicates that most children do attend to the problem.

Inconsistency of response is frequently a problem in any research involving young children. In this experiment, inconsistency was apparent in many, if not most, cases. For example, when asked first whether there were "more" red dots, then "fewer" red dots, subjects often responded "yes" to both questions. Such contradiction, of course, is indicative of failure to understand the problem, as well as a lack of concern for, or awareness of, the logical necessity of answering the questions differently.

As alluded to above, the underlying reason for children's failure to perform well on the class inclusion problem is that they simply do not understand it. An examination of results reveals that even in those experiments which report significant effects of such factors as verbal cueing, or training on the class inclusion problem, children up to at least 10 years of age do not do very well regardless of experimental manipulations. Thus, for example, Sheppard's (1973) training only increased scores from a mean of <1/10 to a mean of approximately 4/10. In other words, no factor has been found

that causes dramatic improvements. It seems appropriate, therefore, to suggest that the concept of class inclusion should be regarded as a formal operational rather than a concrete operational concept; at least until such time as some as yet undiscovered approach to the problem may provide evidence to the contrary.

REFERENCES

Ahr, P. R., & Youniss, J. Reasons for failure on the class inclusion problem. Child Development, 1970, 41, 131-143.

Brainerd, C. J. Training and transfer of transitivity, conservation, and class inclusion of length. Child Development, 1974, 45, 324-334.

Brainerd, C. J. & Kaszor, P. An analysis of two proposed sources of children's class inclusion errors. Developmental Psychology, 1974, 10, 633-643.

Gelman, R. The nature and development of early number concepts. In Hayne W. Reese, (Ed.), Advances in child development and behaviour, vol. 7. New York: Academic Press, 1972.

Inhelder, B., & Piaget, J. The early growth of logic in the child. London: Routledge & Kegan Paul, 1964.

Jennings, J. R. The effect of verbal and pictorial presentation on class-inclusion competence and performance. Psychonomic Science, 1970, 20, 357-358.

Klahr, D., & Wallace, J. G. Class inclusion processes. In S. Farnham-Diggory (Ed.), Information processing in children. New York: Academic Press, 1972.

Kirk, R. E. Experimental design: Procedures for the behavioral sciences. Belmont, California: Brooks/Cole, 1968.

Kofsky, E. A scalogram study of classificatory development. Child Development, 1966, 37, 191-204.

Lawson, G., Baron, J., & Siegel, L. The role of number and length cues in children's quantitative judgments. Child Development, 1974, 45, 731-736.

Piaget, J. The child's conception of number. London: Routledge & Kegan Paul, 1952.

Sheppard, J. L. Conservation of part and whole in the acquisition of class inclusion. Child Development, 1973, 44, 380-383.

Smither, S. J., Smiley, S. S., & Rees, R. The use of perceptual cues for number judgment by young children. Child Development, 1974, 45, 693-699.

Winer, G. A. An analysis of verbal-facilitation of class-inclusion reasoning. Child Development, 1974, 45, 224-227.

Wohlwill, J. F. Responses to class-inclusion questions for verbally and pictorially presented items. Child Development, 1968, 39, 449-465.